

# A QM/MM Study of the Nature of the Entatic State in Plastocyanin

Catherine A. Hurd, Nicholas A. Besley,\* and David Robinson\*

Plastocyanin is a copper containing protein that is involved in the electron transfer process in photosynthetic organisms. The active site of plastocyanin is described as an entatic state whereby its structure represents a compromise between the structures favored by the oxidized and reduced forms. In this study, the nature of the entatic state is investigated through density functional theory-based hybrid quantum mechanics/molecular mechanics (QM/MM) molecular dynamics simulations. The strain energy is computed to be 12.8 kcal/mol and 14.5 kcal/mol for the oxidized and reduced forms of the protein, indicating that the active site has an intermediate

structure. It is shown that the energy gap between the oxidized and reduced forms varies significantly with the fluctuations in the structure of the active site at room temperature. An accurate determination of the reorganization energy requires averaging over conformation and a large region of the protein around the active site to be treated at the quantum mechanical level. © 2016 The Authors. Journal of Computational Chemistry Published by Wiley Periodicals, Inc.

DOI: 10.1002/jcc.24666

## Introduction

Copper containing proteins are naturally occurring proteins involved in many biological processes, including respiration and photosynthesis.<sup>[1–9]</sup> One of the most prominent examples of these proteins is plastocyanin which transfers an electron from the cytochrome *bf* complex to photosystem I in photosynthetic organisms. In the active site of plastocyanin, copper is bound in an approximate trigonal plane to the sulphur of cysteine and two nitrogen atoms of histidine ligands, with a further methionine ligand in an axial position.

During electron transfer copper is reduced from Cu(II) to Cu(I). In its oxidized form, plastocyanin has a  $d^{9}$  electronic configuration where the singly occupied molecular orbital (SOMO) corresponds to an antibonding combination of a copper 3d orbital and a 3p orbital on the sulphur atom of the coordinated cysteine ligand.<sup>[10]</sup> In this state, the protein has a rich spectroscopy, and has been studied by several different spectroscopic techniques including electronic absorption, circular dichroism, magnetic circular dichroism, and x-ray absorption spectroscopies.<sup>[11,12]</sup> Alongside these experimental measurements there have been numerous theoretical studies with a wide variety of different levels of approximation.<sup>[7,13–24]</sup> These studies have led to an understanding of the nature of the observed spectroscopic features and detailed insight into the electronic structure of the active site. For example, the absorption spectrum shows peaks at  $16,700\text{ cm}^{-1}$  and  $12,800\text{ cm}^{-1}$ , which are assigned to a  $\text{Cys}(\pi) \rightarrow \text{SOMO}$  ligand to metal charge-transfer (LMCT) and d ligand field  $\rightarrow \text{SOMO}$  excitations, respectively, the first of which is much more intense than the second.<sup>[7]</sup> At higher energy ( $21,400\text{ cm}^{-1}$ ) there is an additional band which is assigned to the  $\text{Cys}(\sigma) \rightarrow \text{SOMO}$  transition. In plastocyanin this band is very weak, but in the closely related proteins cucumber basic protein and nitrite reductase its intensity increases.<sup>[12,25]</sup> This change can be

associated with changes in the structure of the active site,<sup>[10]</sup> and other work has shown the excitation energy of the intense LMCT band correlate strongly with the copper-cysteine bond length.<sup>[14]</sup> The spectroscopy of the reduced form of the protein contains less information, thus there have been fewer studies that have focused on this form of the protein.<sup>[26–28]</sup>

The structure of the active site plays a key role in the function of the protein. One debated aspect of plastocyanin redox chemistry has been the entatic<sup>[29]</sup> (or reduced rack<sup>[30]</sup>) state, whereby the structure of the active site represents a compromise between the structures favored by the oxidized and reduced forms of the protein. Adopting such an intermediate structure leads to fast electron transfer rates since the reorganization energy will be small. In experiment the reorganization energy has been measured to be 16–28 kcal/mol.<sup>[31–33]</sup> It is thought that the structure for the active site is a consequence of the extended structure of the protein. Indeed, it has been shown with other protein systems that the protein holds transition metal active sites in such a geometry that the conversion between electronic or spin states requires much lower energies than would be the case for the equivalent “model system” of the active site.<sup>[34]</sup>

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

C. A. Hurd, N. A. Besley, D. Robinson  
School of Chemistry, University of Nottingham, University Park, Nottingham,  
NG7 2RD, United Kingdom E-mail: [nick.besley@nottingham.ac.uk](mailto:nick.besley@nottingham.ac.uk) or [david.robinson@nottingham.ac.uk](mailto:david.robinson@nottingham.ac.uk)

Submitted as part of the Simulating Biomolecules: Charlie Brooks' festschrift.  
Contract grant sponsor: Engineering and Physical Sciences Research Council; Contract grant number: EP/I012303; Contract grant sponsor: University of Nottingham for an Impact Acceleration Award and provision of time on the Minerva High-Performance Computing cluster

© 2016 The Authors. Journal of Computational Chemistry Published by Wiley Periodicals, Inc.

This concept of an entatic state was questioned following quantum mechanical calculations of the structure of the active site.<sup>[35–37]</sup> In this work, the structure of the active site was optimized free from the constraints of the protein using density functional theory (DFT) and shown to have a very similar structure to the crystal structure. It was concluded that the active site was not strained to a significant extent in either the oxidized and reduced forms of the protein.<sup>[37]</sup> However, studies of the isolated active site with relatively small model systems will only probe any strain that is localized very near the copper center. These studies initiated further discussion of the copper active site. Further analysis of experimental evidence was found to indicate that the active site of blue copper proteins is constrained for functional advantage with a low reorganization energy for electron transfer.<sup>[38,39]</sup> Deeth studied the oxidized form of a range of blue copper proteins, including plastocyanin, with ligand field molecular mechanics (LFMM) and found an entatic strain energy of about 10 kcal mol<sup>−1</sup>.<sup>[18]</sup> Further hybrid quantum mechanics/molecular mechanics (QM/MM) approaches<sup>[24]</sup> and implicit models of the protein environment<sup>[40]</sup> have been used to obtain optimized geometries of each state.

The studies of the strain energy and the reorganization energy have been largely based on a static view of the structures. It is known that conformational averaging is important in determining the properties of these proteins.<sup>[16,23]</sup> In this article, we develop a more dynamic picture of the electron transfer process. Previous work has studied photoinduced electron transfer dynamics between the oxidized ground state and optically excited LMCT state based on classical molecular dynamics (MD) simulations.<sup>[41]</sup> The parameterization of classical force fields to accurately describe the structure of the active site is problematic, and a previous study has shown that electronic absorption and circular dichroism spectra computed from structures from a classical MD trajectory were in poor agreement with experiment compared with those based on structures from a LFMM-based MD simulation.<sup>[23]</sup> QM/MM approaches have the potential to provide a more accurate and balanced treatment of the oxidized and reduced forms of the protein. However, the size of the QM region in these calculations can be critical and there needs to be a balance between the computational cost of a large QM region and the length of MD simulation that can be performed. Here, we investigate the dynamics and the associated reorganization energy of the oxidized and reduced forms of plastocyanin using a QM/MM approach combining DFT with an all-electron basis set for the active site with the remainder of the protein described by a classical force field. The importance in the choice of the size of the active site treated at the QM level is explored with further calculations with a very large QM region.

## Methods

### Gas-phase quantum chemical calculations

The structure of the gas-phase model active site (see Fig. 1) was optimized with DFT with the  $\omega$ B97X-D functional<sup>[42,43]</sup> and

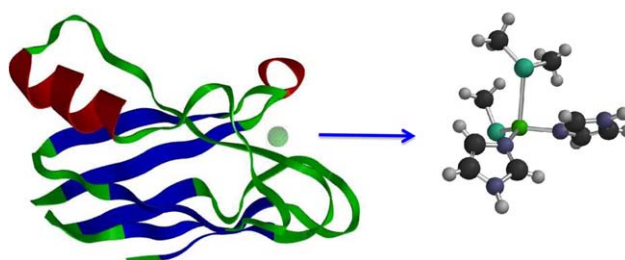


Figure 1. Plastocyanin and the model of the active site used in the QM calculations. [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

6-31G(d)<sup>[44–46]</sup> basis set for both the Cu(I) and Cu(II) oxidation states. The  $\omega$ B97X-D functional incorporates an atom–atom dispersion term and has been shown to be accurate for a wide range of applications including thermochemistry, fundamental gaps, and noncovalent interactions which are relevant for this study.<sup>[43]</sup> Harmonic frequency calculations were performed to ensure the resulting structures corresponded to energy minima.

### QM/MM dynamics

Initially, reduced and oxidized forms of plastocyanin from the *Dryopteris* fern (PDB codes 1KDI and 1KDJ, respectively)<sup>[47]</sup> were minimized and equilibrated using the CHARMM36 protein force-field,<sup>[48,49]</sup> with additional bonding terms for the active site taken from the work of Voth and coworkers,<sup>[41]</sup> the partial atomic charges from the CHARMM36 force-field were retained, while the copper center was assigned the formal charge relevant to the oxidation state. The system was created using the zwitterionic form of plastocyanin and 5496 TIP3P<sup>[50]</sup> water molecules, using a truncated octahedral symmetry. Nineteen K<sup>+</sup> and thirteen Cl<sup>−</sup> ions were added to neutralize the overall charge of the system, giving a physiologically relevant concentration of 0.15 M. Periodic boundaries were employed, using the Particle-mesh Ewald summation method for long-range electrostatics, with a cutoff distance of 12 Å and a pair-list distance of 16 Å. Minimization was performed using the steepest-descent algorithm for 10,000 steps. Equilibration was carried out for a total of 1 ns, with a timestep of 2 fs, with rescaling of velocities every 500 steps. All bonds to hydrogen were constrained using the SHAKE algorithm. This gave an equilibrated protein and water molecules system that provides a suitable starting point for further calculations.

Using the equilibrated system, a QM/MM partitioning was performed, using the active site shown in Figure 1, with hydrogen link-atoms where the bond was broken for His37, Cys87, His90, and Met95, giving a total of 33 QM atoms (identical to the atoms used in the gas-phase calculations). PME was used only for the MM/MM partition, with QM/MM polarization for the QM region. For the case of electron transfer, long-range electrostatics may be important and thus the QM/MM approach may be limited. For both the oxidized and reduced forms, QM/MM minimization was performed until the energy change was less than  $1 \times 10^{-4}$  kcal mol<sup>−1</sup>, before a short period of equilibration was performed (using the same

Table 1. Structural parameters for the oxidized form of plastocyanin.

Geometrical parameters	Active site	QM/MM	Crystal structure <sup>[a]</sup>
<b>Bonds</b>			
Cu—S(cys)	2.18	2.16 (0.05)	2.26
Cu—S(met)	2.36	2.78 (0.16)	2.92
Cu—N(his37)	2.00	1.95 (0.05)	1.93
Cu—N(his90)	1.98	1.91 (0.04)	2.07
<b>Angles</b>			
S(cys)—Cu—S(met)	107.5	102.7 (4.4)	106.7
N(his37)—Cu—N(his90)	97.0	99.5 (3.2)	107.3
S(cys)—Cu—N(his37)	140.1	129.7 (5.3)	126.1
S(cys)—Cu—N(his90)	96.5	122.1 (7.3)	118.3
<b>Dihedral Angles</b>			
C <sub>β</sub> (cys)—S(cys)—Cu—S(met)	41.5	−12.4 (6.8)	−8.9
C <sub>γ</sub> (his37)—N(his37)—Cu—S(met)	22.7	−117.6 (9.0)	−118.6
C <sub>γ</sub> (his90)—N(his90)—Cu—S(met)	88.6	55.3 (9.0)	47.4
<b>Improper Torsions</b>			
Cu—S(cys)—N(his90)—N(his37)	−24.2	−19.1 (5.8)	−25.2
Cu—S(cys)—N(his37)—N(his90)	53.3	22.1 (7.8)	33.1
Cu—N(his90)—N(his37)—S(cys)	−25.5	−15.0 (4.8)	−17.1

Bond lengths in Å and angles in degrees. QM/MM values correspond to average values over the simulation with root mean squared variations given in parenthesis.  
[a] Crystal structure with a resolution of 1.7 Å from Ref. [47]

parameters as above; a total of 2 ps). Production QM/MM dynamics were performed for a total of 40 ps. The ωB97X-D functional and the 6-31G(d) basis set were used throughout. All of the quantum chemical calculations were performed with the Q-Chem software,<sup>[51]</sup> while the QM/MM simulations were performed with the Q-Chem/CHARMM interface,<sup>[52,53]</sup> using the CHARMM36 protein force-field and the TIP3P water model. Further QM calculations were performed with a greatly enlarged QM region that comprised 1144 atoms, chosen by including all residues which are within (or have at least one atom within) a sphere of radius 9 Å of the copper centre, no MM charges were included in these calculations. For these calculations the 6-31G\* basis set was used for the copper and cysteine sulphur atoms, while the 6-31G basis set was used for the remaining atoms. Overall this gave a calculation with nearly 6000 basis functions.

## Results and Discussion

Tables 1 and 2 summarize the values of key geometrical parameters of the active site for both the Cu(II) (oxidized) and Cu(I) (reduced) model compounds. Results are shown for the isolated active site (QM atoms only, no MM charges), average values from the QM/MM simulations and from the crystal structure. Several other computational studies have investigated the structure of the active site of the oxidized and reduced forms. Pavelka and Burda<sup>[40]</sup> found optimized structures of the active site in gas-phase and with implicit solvent models<sup>[39]</sup> to differ significantly from the crystal structure, particularly for the Cu—S(met) bond length for both oxidized and reduced forms of the protein. Relative to the crystal structure, they found this bond length was reduced significantly in the model active sites for the oxidized form while a very large increase was observed for the reduced form to give an unusually large

(essentially not bonded) bond length. The results presented here show a significant decrease in the Cu—S(met) bond length for both the oxidized and reduced forms of the protein with respect to the crystal structures, which is consistent with other studies.<sup>[17,37]</sup> Overall, the structures for the active site in gas-phase show significant differences from the crystal structure for both the bond lengths and dihedral angles to the coordinating ligands.

The average structural parameters from the QM/MM simulations are in closer agreement with the crystal structure, particularly considering some variation between the reported crystal structures and the dynamical nature of the protein in solution is to be expected. The predicted Cu—S(cys) bond length for the oxidized form of 2.16 Å is consistent with the value of 2.15 Å derived from analysis of the absorption spectrum<sup>[20]</sup> and is also agrees with the value of 2.17 Å derived from LFMM-based MD simulations.<sup>[18]</sup> During the QM/MM simulations, the bond lengths typically fluctuate by about 0.05 Å. The exception to this is the Cu—S(met) bond length which has a significantly larger variation during the simulation, and the extent of this variation is found to be greater for the reduced form of the protein. This is consistent with the nature of the Cu—S(met) bond which is known to be a relatively facile bond. Similarly, a large variation in the dihedral angles involving the S(met) ligand are observed. The sensitivity of the Cu—S(met) bond length to the protein environment has been observed in previous studies. Sinnecker and Neese<sup>[17]</sup> found that an isolated model active site was not able to reproduce the long crystal structure Cu—S(met) bond length, however, a much closer value was obtained in QM/MM calculations. Furthermore, Ryde et al. showed this bond length to increase with the value of dielectric constant in calculations with an implicit solvent model.<sup>[37]</sup>

The reduced form exhibits tetrahedral coordination to the copper center, while the oxidized form shows a more distorted geometry. For the oxidized form, it is useful to calculate the geometrical  $\tau_4$  parameter,<sup>[54]</sup> given as

$$\tau_4 = \frac{360^\circ - (\alpha + \beta)}{141^\circ} \quad (1)$$

where  $\alpha$  and  $\beta$  are the two largest bond angles of the copper coordination site and the value of  $\tau_4$  varies between 0 for square-planar and 1 for tetrahedral geometries. For the isolated active site values of  $\tau_4 = 0.57$  and  $\tau_4 = 0.91$  are obtained for the oxidized and reduced forms. This shows that the reduced form is close to tetrahedral while the oxidized form is intermediate between square-planar and tetrahedral; this is typical for 4-coordinate Cu(II) systems. For the average structures from the QM/MM simulations the value of  $\tau_4$  for the oxidized form increases to 0.88, while the reduced form has a value close to 1. Consequently, the distortion from tetrahedral evident in the model gas phase structure of the oxidized form is decreased significantly.

The strain energy provides a measure of the extent to which the protein holds the active site in an energetically nonoptimum structure, and is calculated from the difference between

Table 2. Structural parameters for the reduced form of plastocyanin.

Geometrical parameters	Active site	QM/MM	Crystal structure <sup>[a]</sup>
<b>Bonds</b>			
Cu—S(cys)	2.32	2.16 (0.06)	2.21
Cu—S(met)	2.30	2.63 (0.29)	2.91
Cu—N(his37)	2.00	1.97 (0.08)	1.95
Cu—N(his90)	1.98	1.92 (0.08)	2.10
<b>Angles</b>			
S(cys)—Cu—S(met)	113.8	107.1 (6.8)	108.2
N(his37)—Cu—N(his90)	118.3	100.6 (5.8)	104.6
S(cys)—Cu—N(his37)	103.9	129.4 (8.4)	130.1
S(cys)—Cu—N(his90)	108.8	112.8 (10.2)	117.3
<b>Dihedral Angles</b>			
C <sub>β</sub> (cys)—S(cys)—Cu—S(met)	−40.1	−6.1 (10.2)	−3.8
C <sub>γ</sub> (his37)—N(his37)—Cu—S(met)	−110.2	−131.5 (10.2)	−114.6
C <sub>γ</sub> (his90)—N(his90)—Cu—S(met)	134.5	54.2 (11.6)	51.7
<b>Improper Torsions</b>			
Cu—S(cys)—N(his90)—N(his37)	−31.3	−25.2 (8.4)	−17.6
Cu—S(cys)—N(his37)—N(his90)	29.7	33.1 (9.6)	22.8
Cu—N(his90)—N(his37)—S(cys)	−40.6	−22.6 (6.3)	−15.9

Bond lengths in Å and angles in degrees. QM/MM values correspond to average values over the simulation with root mean squared variations given in parenthesis.  
[a] Crystal structure with a resolution of 1.8 Å from Ref. 47.

the energy of the ensemble averaged QM/MM active site structure in the gas-phase (i.e., no MM charges included) and the energy of the corresponding optimized active-site gas-phase structure. This gives values of 12.8 kcal/mol and 14.5 kcal/mol for the oxidized and reduced forms of the protein, respectively. The value for the oxidized form is close to the value of ~10 kcal/mol determined from LFMM calculations.<sup>[18]</sup> Here, it is shown that the value for the reduced form has a similar magnitude to that of the oxidized form. In both cases, there is significant strain on the coordination center. These values are consistent with the premise of an entatic state where the geometry of the active site is intermediate between the optimal geometries of the two states.

Figure 2 shows the variation in  $\Delta E$ , where  $\Delta E = E_{\text{ox}} - E_{\text{red}}$ , during the QM/MM simulations for the oxidized and reduced forms of the protein. This corresponds to the energy change between the two states without allowing for any relaxation of the structure. For both trajectories there is significant variation in  $\Delta E$ . For the reduced trajectory, the average value of  $\Delta E$  is 82.5 kcal/mol with a standard deviation of 11.2 kcal/mol. The oxidized trajectory shows a surprising amount of deviation in  $\Delta E$ , which is initially positive and then falls rapidly, where it varies predominantly between −10 kcal/mol and −40 kcal/mol. The initial rapid change in  $\Delta E$  at the start of the simulation is likely to reflect that the protein was still equilibrating during this period. Focusing on the simulation from 3 ps onwards gives an average value for  $\Delta E$  of −29.2 kcal/mol with a standard deviation of 9.0 kcal/mol. This variation in  $\Delta E$  suggests that the view of the protein being held in a single entatic state is not the complete picture since an ensemble of structures with significantly different  $\Delta E$  will be sampled.

The reorganization energy is a key component in Marcus theory where the rate of electron transfer is given by<sup>[55]</sup>

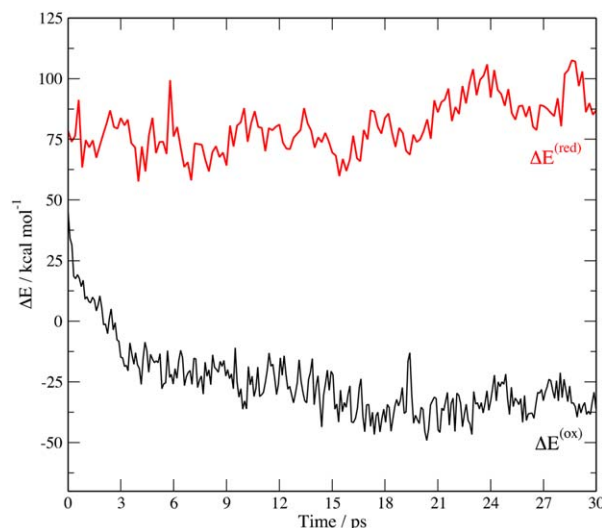


Figure 2. Variation in  $\Delta E$  ( $E_{\text{ox}} - E_{\text{red}}$ ) during the QM/MM molecular dynamics simulations for the oxidized and reduced forms of the protein. [Color figure can be viewed at wileyonlinelibrary.com]

$$k_{\text{ET}} = \frac{|V|^2}{\hbar} \sqrt{\frac{\pi}{\lambda k_{\text{B}} T}} \exp \left[ -\frac{(\lambda + \Delta G)^2}{4\lambda k_{\text{B}} T} \right] \quad (2)$$

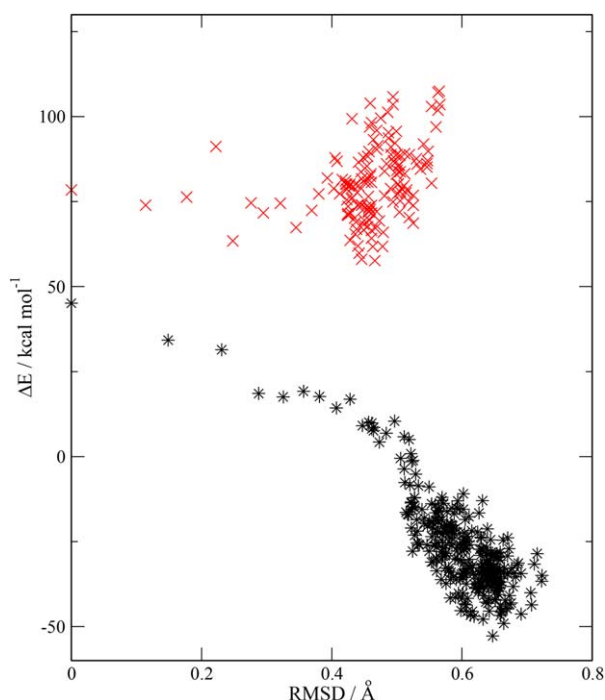
where  $\Delta G$  is often termed the driving force of the reaction and is the change in free energy between the final and initial states and  $V$  is the diabatic electron coupling matrix element between the two states. The total reorganization energy,  $\lambda$ , can be calculated from the average of the instantaneous difference in energy between the oxidized and reduced forms ( $\Delta E$ ) from both the oxidized trajectory and the reduced trajectory. Within the linear response approximation the reorganization energy is given by<sup>[56]</sup>:

$$\lambda = \frac{1}{2} (\langle \Delta E \rangle_{\text{red}} - \langle \Delta E \rangle_{\text{ox}}) \quad (3)$$

where  $\langle \Delta E \rangle_{\text{red}}$  and  $\langle \Delta E \rangle_{\text{ox}}$  are the average values of  $\Delta E$  from the reduced and oxidized trajectories, respectively. The total reorganization energy has been measured to lie between 16 kcal/mol and 28 kcal/mol,<sup>[31–33]</sup> and estimated to be about 16 kcal/mol and 23 kcal/mol from DFT and MP2 calculations.<sup>[37]</sup> The present results for the overall average reorganization energy, from eq. (3), find  $\lambda = 55.9$  kcal/mol, which is significantly larger than the experimental value, predominantly owing to the high value for  $\langle \Delta E \rangle_{\text{red}}$ . The accuracy of the computed value of  $\lambda$  can depend on a number of factors including applicability of the linear response approximation, the completeness of the sampling and the accuracy of the QM/MM method. The computational cost of QM/MM approaches makes it challenging to achieve a thorough sampling of configurations, and we explore the accuracy of the QM/MM model in more detail later.

Figure 3 illustrates the correlation between the root mean squared deviation (RMSD) in the structure relative to the structure at  $t=0$  and the calculated  $\Delta E$  for the oxidized and reduced trajectories. For the oxidized form, the plot suggests a

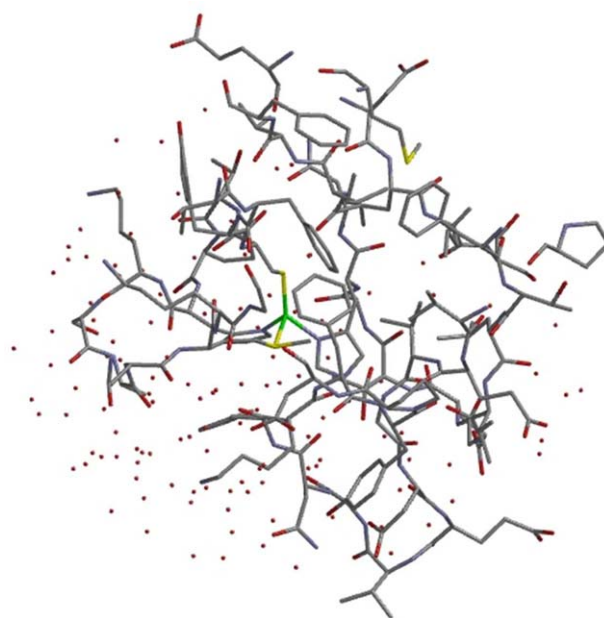




**Figure 3.** Correlation between  $\Delta E$  and the RMSD of the protein structure relative to the structure at  $t = 0$  for the QM/MM simulations. Oxidized form \*, reduced form  $\times$ . [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

significant degree of correlation between the structure and the calculated  $\Delta E$ . There is also a large amount of change in the RMSD for the structures in the early part of the simulation (those with positive  $\Delta E$ ) confirming that the protein was still undergoing some equilibration in this period. For the reduced form there is less variation in the RMSD during the simulation, and correspondingly less variation in  $\Delta E$ .

Assuming that the electronic coupling matrix elements do vary significantly with the fluctuations in the structure,<sup>[57]</sup> the most efficient electron transfer will occur at the point when the reorganization energy is smallest, provided that this structure is thermally accessible. The oxidized trajectory reveals some interesting structures where the energy difference between the oxidized and reduced forms of the protein is close to zero. The value with smallest magnitude was  $-0.5$  kcal/mol at 2.6 ps. The corresponding active site is distorted with  $\tau_4 = 0.76$ , and if such a structure was thermally accessible it would result in very fast rates of electron transfer. To explore the energetics of these structures further, additional calculations with a larger QM region describing the active site have been performed for the structures corresponding to the times of 2.6 ps and 24.4 ps. These structures were chosen based on their associated  $\Delta E$ , with the structure at 2.6 ps corresponding to the minimum value and the structure at 24.4 ps close to the average value during the part of the simulation where the structure had stabilized. The larger QM region is shown in Figure 4, and represents the largest QM region that was tractable with our computational resources. This will account more accurately for structural effects further from the active site, and also has the effect of removing the point charges of the



**Figure 4.** Enlarged active site region treated at the QM level, the copper atom is shown in green and the hydrogen atoms are omitted for clarity. [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

MM region from close to the copper center giving insight into the accuracy of a QM/MM partitioning in these systems. With the larger QM region the computed  $\Delta E$  change from  $-0.5$  kcal/mol to  $-7.0$  kcal/mol at 2.6 ps and  $-25.0$  kcal/mol to  $-17.7$  kcal/mol at 24.4 ps. This shows that the smaller QM active site calculations correctly predict that the structure at 2.6 ps has the smaller reorganization energy. However, while the variation in the reorganization energy remains significant, it suggests that the smaller QM region tends to exaggerate these fluctuations. This is likely to be associated with having a larger number of classical point charges very close to the copper active site center.

## Conclusions

QM/MM molecular dynamics simulations have been performed for the oxidized and reduced forms of plastocyanin. The structural parameters describing the active site determined through averaging the geometries through the MD simulation are close to those of the crystal structure and show some significant differences to those of the optimized isolated gas phase active site. The strain energy for the active site is computed to be 12.8 kcal/mol for the oxidized form, in agreement with previous work.<sup>[18]</sup> For the reduced form the strain energy is found to be 14.5 kcal/mol, which is comparable to the value for the oxidized form. This is consistent with the concept of an entatic state where the protein scaffold maintains a structure that is nonoptimal for either oxidized or reduced forms of the protein.

There is considerable variation in the computed instantaneous energy difference between the oxidized and reduced forms of the protein arising from the thermal fluctuations in the structure at room temperature. This necessitates some

averaging over conformation in determining the reorganization energy, and means that the view of the protein constraining the active site in a structure that is optimal for electron transfer does not provide the complete picture of the electron transfer process. The computed value for  $\lambda$  is higher than values estimated from experiment, although the simulations do sample structures that have a much lower value for  $\lambda$ . This may indicate more extensive sampling is required. The evidence of the calculations presented here is that using a QM region that includes the typical active site treated in QM calculations reproduces the trends in the energy gap between oxidized and reduced forms, but tends to overestimate the fluctuations. This illustrates the challenge of modeling the electron transfer in these systems within a QM/MM framework. Treating a very large region of the protein surrounding the active site at the QM level is desirable for a reliable estimate of the reorganization energy. However, these calculations are very computationally demanding and are not practical for the study of many conformations or the use in molecular dynamics simulations. One solution to this problem might be to use more approximate methods, such as LFMM, provided they give a balanced treatment of both oxidized and reduced forms of the protein.

**Keywords:** QM/MM • electron transfer • plastocyanin

How to cite this article: C. A. Hurd, N. A. Besley, D. Robinson. *J. Comput. Chem.* **2017**, 38, 1431–1437. DOI: 10.1002/jcc.24666

- [1] P. M. Colman, H. C. Freeman, J. M. Guss, M. Murata, V. A. Norris, J. A. M. Ramshaw, M. P. Venkatappa, *Nature* **1978**, 272, 319.
- [2] T. Inoue, H. Sugawara, S. Hamanaka, H. Tsukui, E. Suzuki, T. Kohzuma, Y. Kai, *Biochemistry* **1999**, 38, 6063.
- [3] N. Shibata, T. Inoue, C. Nagano, N. Nishio, T. Kohzuma, K. Onodera, F. Yoshizaki, Y. Sugimura, Y. Kai, *J. Biol. Chem.* **1999**, 274, 4225.
- [4] Y. F. Xue, M. Okvist, O. Hansson, S. Young, *Prot. Sci.* **1998**, 7, 2099.
- [5] A. J. Wachtters, D. Cascio, M. K. Choukair, D. Rice, S. Merchant, T. O. Yeates, *Biochemistry* **1993**, 32, 10560.
- [6] E. I. Solomon, K. W. Penfield, A. A. Gewirth, M. D. Lowery, S. E. Shadle, J. A. Guckert, L. B. LaCroix, *Inorg. Chim. Acta* **1996**, 243, 67.
- [7] A. A. Gewirth, E. I. Solomon, *J. Am. Chem. Soc.* **1988**, 110, 3811.
- [8] K. W. Penfield, A. A. Gewirth, E. I. Solomon, *J. Am. Chem. Soc.* **1985**, 107, 4519.
- [9] K. W. Penfield, R. R. Gray, R. S. Himmelwright, N. C. Eickman, V. A. Norris, H. C. Freeman, E. I. Solomon, *J. Am. Chem. Soc.* **1981**, 103, 4382.
- [10] E. I. Solomon, R. K. Szilagyi, S. DeBeer George, L. Basumallick, *Chem. Rev.* **2004**, 104, 419.
- [11] D. W. Randall, S. DeBeer George, B. Hedman, K. O. Hodgson, K. Fujisawa, E. I. Solomon, *J. Am. Chem. Soc.* **2000**, 122, 11620.
- [12] L. B. LaCroix, S. E. Shadle, Y. Wang, B. A. Averill, B. Hedman, K. O. Hodgson, E. I. Solomon, *J. Am. Chem. Soc.* **1996**, 118, 7755.
- [13] S. Larsson, A. Broo, L. Sjoelin, *J. Phys. Chem.* **1995**, 99, 4860.
- [14] K. Pierloot, J. O. A. De Kerpel, U. Ryde, B. O. Roos, *J. Am. Chem. Soc.* **1997**, 119, 218.
- [15] K. Pierloot, J. O. A. De Kerpel, U. Ryde, M. H. M. Olsson, B. O. Roos, *J. Am. Chem. Soc.* **1998**, 120, 13156.
- [16] M. H. M. Olsson, G. Hong, A. Warshel, *J. Am. Chem. Soc.* **2003**, 125, 5025.
- [17] S. Sinnecker, F. Neese, *J. Comput. Chem.* **2006**, 27, 1463.
- [18] R. J. Deeth, *Inorg. Chem.* **2007**, 46, 4492.
- [19] S. Vancoillie, K. Pierloot, *J. Phys. Chem. A* **2008**, 112, 4011.
- [20] D. Robinson, N. A. Besley, *Phys. Chem. Chem. Phys.* **2010**, 12, 9667.
- [21] N. A. Besley, D. Robinson, *Faraday Discuss.* **2011**, 148, 55.
- [22] A. Monari, T. Very, J. –L. Rivail, X. Assfeld, *Comput. Theor. Chem.* **2012**, 990, 119.
- [23] H. Do, R. J. Deeth, N. A. Besley, *J. Phys. Chem. B* **2013**, 117, 8105.
- [24] H. Y. Jin, P. Goyal, A. K. Das, M. Gaus, M. Meuwly, Q. Cui, *J. Phys. Chem. B* **2016**, 120, 1894.
- [25] L. B. LaCroix, D. W. Randall, A. M. Nersissian, C. W. G. Hoitink, G. W. Canters, J. S. Valentine, E. I. Solomon, *J. Am. Chem. Soc.* **1998**, 120, 9621.
- [26] J. A. Guckert, M. D. Lowery, E. I. Solomon, *J. Am. Chem. Soc.* **1995**, 117, 2817.
- [27] L. Shan Kau, D. J. Spira-Solomon, J. E. Penner-Hahn, K. O. Hodgson, E. I. Solomon, *J. Am. Chem. Soc.* **1987**, 109, 6433.
- [28] H. Do, N. A. Besley, *J. Phys. Chem. A* **2012**, 116, 8507.
- [29] B. L. Vallee, R. J. P. Williams, *Proc. Natl. Acad. Sci. USA* **1968**, 59, 498.
- [30] B. G. Malmstrom, *Eur. J. Biochem.* **1994**, 223, 711.
- [31] A. J. Di Bilio, M. G. Hill, N. Bonander, B. G. Karlsson, R. M. Villahermosa, B. G. Malmström, J. R. Winkler, H. B. Gray, *J. Am. Chem. Soc.* **1997**, 119, 9921.
- [32] L. K. Skov, *J. Am. Chem. Soc.* **1998**, 120, 1102.
- [33] K. Sigfridsson, M. Ejdebäck, M. Sundahl, Ö. Hansson, *Biochem. Biophys.* **1998**, 351, 197.
- [34] P. Jayapal, D. Robinson, M. Sundararajan, I. H. Hillier, J. J. W. McDouall, *Phys. Chem. Chem. Phys.* **2008**, 10, 1734.
- [35] U. Ryde, M. H. M. Olsson, K. Pierloot, B. O. Roos, *J. Mol. Biol.* **1996**, 261, 586.
- [36] M. H. M. Olsson, U. Ryde, B. O. Roos, K. Pierloot, *J. Biol. Inorg. Chem.* **1998**, 3, 109.
- [37] U. Ryde, M. H. M. Olsson, B. O. Roos, A. C. Borin, *Theor. Chem. Acc.* **2001**, 105, 452.
- [38] H. B. Gray, B. G. Malmstrom, R. J. P. Williams, *J. Biol. Inorg. Chem.* **2000**, 5, 551.
- [39] U. Ryde, M. H. M. Olsson, B. O. Roos, J. O. A. De Kerpel, K. Pierloot, *J. Biol. Inorg. Chem.* **2000**, 5, 565.
- [40] M. Pavelka, J. V. Burda, *Mol. Phys.* **2008**, 106, 2733.
- [41] L. W. Ungar, N. F. Scherer, G. A. Voth, *Biophys. J.* **1997**, 72, 5.
- [42] J.-D. Chai, M. Head-Gordon, *J. Chem. Phys.* **2008**, 128, 084106.
- [43] Y.-S. Lin, G.-D. Li, S.-P. Mao, J.-D. Chai, *J. Chem. Theory Comput.* **2013**, 9, 263.
- [44] P. C. Hariharan, J. A. Pople, *Theor. Chim. Acta* **1973**, 28, 213.
- [45] M. M. Francl, W. J. Pietro, W. J. Hehre, J. S. Binkley, M. S. Gordon, D. J. DeFrees, J. A. Pople, *J. Chem. Phys.* **1982**, 77, 3654.
- [46] V. Rassolov, J. A. Pople, M. Ratner, T. L. Windus, *J. Chem. Phys.* **1998**, 109, 1223.
- [47] T. Kohzuma, T. Inoue, F. Yoshizaki, Y. Sasakawa, K. Onodera, S. Nagatomo, T. Kitagawa, S. Uzawa, Y. Isobe, Y. Sugimura, M. Gotowda, Y. Kai, *J. Biol. Chem.* **1999**, 274, 11817.
- [48] R. B. Best, X. Zhu, J. Shim, P. E. M. Lopes, J. Mittal, M. Feig, A. D. MacKerell, Jr., *J. Chem. Theory Comput.* **2012**, 8, 3257.
- [49] A. D. MacKerell, Jr., D. Bashford, M. Bellott, R. L. Dunbrack, Jr., J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, K. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, III, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, M. Karplus, *J. Phys. Chem. B* **1998**, 102, 3586.
- [50] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *J. Chem. Phys.* **1983**, 79, 926.
- [51] Y. Shao, Z. Gan, E. Epifanovsky, A. T. B. Gilbert, M. Wormit, J. Kussmann, A. W. Lange, A. Behn, J. Deng, X. Feng, D. Ghosh, M. Goldey, P. R. Horn, L. D. Jacobson, I. Kaliman, R. Z. Khaliullin, T. Kus, A. Landau, J. Liu, E. I. Proynov, Y. Min Rhee, R. M. Richard, M. A. Rohrdanz, R. P. Steele, E. J. Sundstrom, H. L. Woodcock, III, P. M. Zimmerman, D. Zuev, B. Albrecht, E. Alguire, B. Austin, G. J. O. Beran, Y. A. Bernard, E. Berquist, K. Brandhorst, K. B. Bravaya, S. T. Brown, D. Casanova, C. M. Chang, Y. Chen, S. H. Chien, K. D. Closser, D. L. Crittenden, M. Diedenhofen, R. A. DiStasio, Jr., H. Do, A. D. Dutoi, R. G. Edgar, S. Fatehi, L. Fusti-Molnar, A. Ghysels, A. Golubeva-Zadorozhnaya, J. Gomes, M. W. D. Hanson-Heine, P. H. P. Harbach, A. W. Hauser, E. G. Hohenstein, Z. C. Holden, T. C. Jagau, H. Ji, B. Kaduk, K. Khistyayev, J. Kim, J. Kim, R. A. King, P. Klunzinger, D. Kosenkov, T. Kowalczyk, C. M.

- Krauter, K. Un Lao, A. D. Laurent, K. V. Lawler, S. V. Levchenko, C. Y. Lin, F. Liu, E. Livshits, R. C. Lochan, A. Luenser, P. Manohar, S. F. Manzer, S. P. Mao, N. Mardirossian, A. V. Marenich, S. A. Maurer, N. J. Mayhall, E. Neuscamman, C. M. Oana, R. Olivares-Amay, D. P. O'Neill, J. A. Parkhill, T. M. Perrine, R. Peverati, A. Prociuk, D. R. Rehn, E. Rosta, N. J. Russ, S. M. Sharada, S. Sharma, D. W. Small, A. Sodt, T. Stein, D. Stuck, Y. C. Su, A. J. W. Thom, T. Tsuchimochi, V. Vanovschi, L. Vogt, O. Vydrov, T. Wang, M. A. Watson, J. Wenzel, A. White, C. F. Williams, J. Yang, S. Yeganeh, S. R. Yost, Z. Q. You, I. Y. Zhang, X. Zhang, Y. Zhao, B. R. Brooks, G. K. L. Chan, D. M. Chipman, C. J. Cramer, W. A. Goddard, III, M. S. Gordon, W. J. Hehre, A. Klamt, H. F. Schaefer, III, M. W. Schmidt, C. D. Sherrill, D. G. Truhlar, A. Warshel, X. Xu, A. Aspuru-Guzik, R. Baer, A. T. Bell, N. A. Besley, J. D. Chai, A. Dreuw, B. D. Dunietz, T. R. Furlani, S. R. Gwaltney, C. P. Hsu, Y. Jung, J. Kong, D. S. Lambrecht, W. Liang, C. Ochsenfeld, V. A. Rassolov, L. V. Slipchenko, J. E. Subotnik, T. Van Voorhis, J. M. Herbert, A. I. Krylov, P. M. W. Gill, M. Head-Gordon, *Mol. Phys.* **2015**, *113*, 184.
- [52] H. Lee Woodcock, M. Hodosceck, A. T. B. Gilbert, P. M. W. Gill, H. F. Schaefer, B. R. Brooks, *J. Comput. Chem.* **2007**, *28*, 1485.
- [53] B. R. Brooks, C. L. Brooks, III, A. D. MacKerell, Jr., L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York, M. Karplus, *J. Comput. Chem.* **2009**, *30*, 1545.
- [54] L. Yang, D. R. Powell, R. P. Houser, *Dalton Trans.* **2007**, *9*, 955.
- [55] R. A. Marcus, *J. Chem. Phys.* **1956**, *24*, 966.
- [56] I. Muegge, P. X. Qi, A. J. Wand, Z. T. Chu, A. Warshel, *J. Phys. Chem. B* **1997**, *101*, 825.
- [57] J. E. Subotnik, J. Vura-Weis, A. J. Sodt, M. A. Ratner, *J. Phys. Chem. A* **2010**, *114*, 8665.

---

Received: 1 August 2016

Revised: 23 October 2016

Accepted: 27 October 2016

Published online on 14 November 2016